Real-Time Multiple Event Detection and Classification Using Moving Window PCA

Rafferty, M., Liu, X., Laverty, D. M., & McLoone, S. (2016). Real-Time Multiple Event Detection and Classification Using Moving Window PCA. *IEEE Transactions on Smart Grid*, *7*(5), 2537-2548. https://doi.org/10.1109/TSG.2016.2559444

Published in:
IEEE Transactions on Smart Grid

Document Version:
Peer reviewed version

Real-Time Multiple Event Detection and Classification Using Moving Window PCA

M. Rafferty, *Student Member, IEEE,* X. Liu, *Member, IEEE,* D. Laverty, *Member, IEEE* S. McLoone, *Senior Member, IEEE*

***Abstract*—This paper proposes a method for the detection and classification of multiple events in an electrical power system in real-time, namely; islanding, high frequency events (loss of load) and low frequency events (loss of generation). This method is based on principal component analysis of frequency measurements and employs a moving window approach to combat the time-varying nature of power systems, thereby increasing overall situational awareness of the power system. Numerical case studies using both real data, collected from the UK power system, and simulated case studies, constructed using DigSilent PowerFactory, for islanding events, as well as both loss of load and generation dip events, are used to demonstrate the reliability of the proposed method.**

***Index Terms*—Detection and classification, multiple cascading events, principal component analysis, PMU data, real-time, wide-area monitoring.**

I. Introduction

In many countries the rapid growth in energy demand, fuelled by population growth and industrialisation, has put enormous stress on electrical power system infrastructure and presented major challenges to power system operators in terms of their ability to provide a stable and sustainable supply of electricity to meet consumer demands. In particular, the need to operate power systems close to their limits [1] to meet demand has increased their sensitivity to faults, in extreme circumstances resulting in the occurrence of large scale blackouts. One such example is the North East Blackout of 2003 where 50 million people were affected, and 63 GW of load was interrupted [2]. The major factors contributing to this blackout were a lack of adequate situational awareness capability and a failure to provide effective real-time diagnostic support [2].

To combat these occurrences, many countries have invested heavily in smart grid technologies, with an emphasis placed on PMU-based Wide Area Monitoring Systems (WAMS), of which event detection and analysis techniques are key parts [3]. PMUs are an accurate and advanced instrument for conducting time synchronised measurements of system conditions and parameters [4]. With more PMUs, or other Intelligent Electronic Devices (IEDs) with PMU functionality coming online, wide area monitoring of power system dynamics becomes possible.

However, the increase in data that comes with the increasing proliferation of PMUs leaves system operators facing a major challenge; namely, how to efficiently process and extract useful information from this data, as it is not feasible for human operators to process the raw data manually. This motivates the need for multivariate statistical techniques for intelligent data analysis to assist real-time decision making.

Principal Component Analysis (PCA), originally developed in 1901, is one of the most popular multivariate statistical techniques for dimensionality reduction and has been widely used in various fields [5], for example, process monitoring in the chemical industry [6] and facial recognition in image processing [7]. It transforms highly correlated variables in a dataset, into a reduced set of uncorrelated variables that best summarise the information in the data. One attractive feature of PCA is that it can be computed efficiently, one component at a time, in a recursive manner. This allows fast computation when only a small number of components are needed, which is generally the case when there is significant redundancy in datasets. PCA has received increasing attention in the power system domain, with applications including fault detection [8], extraction and classification of fault features [9], and fault localisation [10].

Since sudden, rapid changes in system frequency are indicative of undesirable imbalance between load and generation on a power system, frequency is the ideal signal to monitor and analyse for event detection at a system level [11]. Much research has been published on this topic, including for example [12], [13], [14] and [15]. Several PCA-based wide area frequency monitoring approaches have also been proposed [8], [16] and [17], with [18] proposing a monitoring scheme that utilises both frequency and voltage data.

In a recent paper by Barocio et al. [3], a PCA-based statistical monitoring approach was introduced for event detection on simulated voltage data, which relies on two statistics, a Hotelling's T^2 and a Q statistic to determine if the power system is in an abnormal situation. In [18], it was also demonstrated that PCA is a very powerful tool for dimensionality reduction and early event detection. In contrast to [3], the monitoring index in [18] is essentially based on the error between the predicted and the measured values.

Despite the success of PCA-based event detection in previous work, to date the problem of classifying multiple events in real-time has not been fully addressed. Similar to the work of Barocio et al. [3], our previous paper [8] employed a PCA-based statistical monitoring framework for islanding detection. In particular, [8] focuses on the geometric interpretation of the Hotelling's T^2 and a Q statistics, and demonstrates that using frequency data the Q statistic detects islanding events (i.e. where a distributed generator continues to energise local

M. Rafferty, X. Liu, D. Laverty and S. McLoone are with the School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast, BT9 5AH, U.K. e-mail: x.liu@qub.ac.uk

loads after isolation from the main power system), while the T^2 statistic detects generation mismatch events.

Guo et al. [16] proposed a recursive PCA algorithm to combat the time varying nature of power systems in the hope of reducing the occurrence of false alarms compared to a traditional PCA method for islanding detection. Their algorithm updates the PCA model correlation matrix recursively as new data samples become available. However, a weakness of the approach taken is that the size of the updated model grows with the number of data samples, which is constantly increasing. In addition, control limits are not adapted to reflect the normal time variation in the power system.

For enhanced situational awareness it is desirable to be able to classify generation mismatch events as either loss of load events (leading to an increase in frequency) and generation dip events (leading to a decrease in frequency). The PCA based approaches in [8] and [16] do not have the ability to distinguish between (classify) these types of event.

Phillips and Overbye [19] and Gaouda et al. [20] demonstrated how to classify events using pattern recognition techniques but with the main focus on voltage not frequency. Classification is obtained by comparing voltage features (at times when the standard deviation is greater than a set threshold) that most closely match features in reference event data.

In [21] Zheng and Craven develop a Support Vector Machine (SVM) based classifier to distinguish between different events in the power system using frequency data. They argue that this machine learning based approach ensures security and reliability in smart grid infrastructure due to its ability to perform intelligent, robust and fast detection and classification of power system disturbances. However, one main constraint of this method, as well as other SVM methods, is their computational inefficiency. In addition, similarly to [19], the approach requires faulty data in the initial training phase and therefore may not have the ability to classify faults that have not been encountered by the system before.

Bykhovsky and Chow [15] successfully used a clustering method to distinguish between different events in a power system using recorded frequency data. This was achieved by grouping together similar events based on the magnitude of frequency change and the rate of change of frequency from historical data.

Wang et al. [14], propose an unmixing method using sparse coding techniques, for the detection and recognition of multiple events in the power system. They propose that each multiple event is made up of root events, and using a clustering method, they group together events with similar characteristics to learn the patterns of the root events. Sparse coding techniques enable the contribution of each root event to the observed mixed event to be determined.

The importance of generation-load mismatch detection and analysis has been emphasized in Gardner and Liu [22] who propose an approach based on pair-wise comparison of frequency measurements from multiple locations. However, their approach is difficult to implement in practice when there a large number of PMUs and provides a delayed detection of events.

Building on our previous work in [8], this paper presents a novel method to address the challenging problem of distinguishing between multiple cascading events, namely islanding, high frequency and low frequency events on a real-time basis, with a moving window PCA approach adopted to compensate for the time-varying characteristic of the power system.

To summarize, the main contributions of the paper are: (1) the use of moving window PCA (rather than static or recursive PCA) to provide thresholds for event detection that adapt to the time varying behaviour of a power system as reflected in the frequency data; (2) the development of a method to automatically distinguish between high and low frequency events for both islanded and non-islanded power system scenarios; (3) through the combination of (1) and (2) the development of a methodology for real-time detection and classification of high and low frequency events, and islanding, under complex scenarios such as cascading events, thus allowing the system operator improved wide-area situational awareness of the power system; (4) validation of the methodology using real case studies, recorded from the UK power system, and simulated case studies, using DigSilent PowerFactory.

The rest of the paper is organised as follows: Section II introduces our proposed Moving Window PCA (MW-PCA) based Multiple Event Detection and Classification (MEDC) methodology. Its performance is then demonstrated for simulated and real power system frequency events case studies in Sections III and IV, respectively. The effectiveness of the MEDC monitoring scheme is discussed in Section V and finally, conclusions are presented in Section VI.

II. METHODOLOGY

A. Principal Component Analysis

The most common application of PCA is reducing the dimensionality of datasets, typically consisting of large numbers of correlated variables, with minimal information loss [5] in order to reveal any simplified structures that may underline them. This is accomplished by transforming the original variables in the data set to a new set of variables, called Principal Components (PCs); these are uncorrelated and ordered so that the first few preserve most of the variation that was present in the original set of variables [5]. For a more detailed description of PCA-based statistical monitoring the interested reader is referred to [5], [23].

Denoting the set of m raw PMU frequency measurements at the i-th sample instant as $\boldsymbol{f}_i \in \Re^{1\times m}$, a data matrix $\boldsymbol{F} \in \Re^{n\times m}$ can be constructed with each row representing a sample (n = number of samples). After normalisation, such that each column has zero-mean and unit variance, PCA can decompose $\boldsymbol{F}$ into a score matrix, $\boldsymbol{T} \in \Re^{n\times k}$ and a loading matrix, $\boldsymbol{P} \in \Re^{m\times k}$, where k is the number of retained principal components, and $k \leq m$ [24] [23]:

$$\boldsymbol{F} = \boldsymbol{t}_1\boldsymbol{p}_1^T + \boldsymbol{t}_2\boldsymbol{p}_2^T + ... + \boldsymbol{t}_k\boldsymbol{p}_k^T = \boldsymbol{T}\boldsymbol{P}^T + \boldsymbol{E} \qquad (1)$$

with $\boldsymbol{t}$, $\boldsymbol{p}$ and $\boldsymbol{E}$ representing the score vectors, the loading vectors and residual matrix, respectively. As discussed in [8], PCA allows the construction of two statistics, Hotelling's T^2 and Q statistics, which can be used to detect abnormal system

behaviour. The T^2 statistic, a scaled 2-norm of an original sample, $\boldsymbol{f}$, from its mean [23] measures significant variation of the recorded data, and is defined as [24] :

$$T^2 = (\boldsymbol{f} - \bar{\boldsymbol{f}})\boldsymbol{P}\Lambda^{-1}\boldsymbol{P}^T(\boldsymbol{f} - \bar{\boldsymbol{f}})^T \tag{2}$$

where Λ is a diagonal matrix containing the k eigenvalues of the covariance matrix of $\boldsymbol{F}$ and $\bar{\boldsymbol{f}}$ is the mean of the frequency data.

The Q statistic, or squared prediction error (SPE), is a squared 2-norm measuring the deviation of the observations from the lower-dimensional PCA representation [24] and is defined as [23]:

$$Q = \boldsymbol{e}^T\boldsymbol{e}, \qquad \boldsymbol{e} = [\boldsymbol{I} - \boldsymbol{P}\boldsymbol{P}^T]\boldsymbol{f} \tag{3}$$

where $\boldsymbol{e}$ is a residual vector and $\boldsymbol{I}$ is an identity matrix. Under normal operation, frequency variation in a power system is approximately normally distributed [8], hence it follows that T^2 and Q can be approximated as central χ^2 distributions [10], that is:

$$T^2 \quad \& \quad Q : g.\chi^2(h) \tag{4}$$

where $g = p^2/2\mu$,$h = 2\mu^2/p^2$, with p^2 and μ as the sample variance and mean of the statistics, respectively, and hence the confidence limits for each statistic, denoted as T^2_α and Q_α, can be computed [17].

The geometric interpretation of PCA introduced in our previous work, [8], shows how PCA has the ability to detect faults. When the system is considered in normal operation the calculated values of these statistics should be quite low (mostly falling below their associated confidence limits). If a calculated statistic violates its confidence limit for a number of samples then this implies abnormal system behaviour and hence the occurrence of an event on the power system.

B. Moving Window Principal Component Analysis

The continuously changing nature of frequency in a power system, determined and controlled by the real-time balance between demand and generation, is such that if demand is greater than generation the system frequency falls below the target mains frequency and increases above it if generation is greater than demand [11]. To track these normal changes in frequency an adaptive PCA method is needed which can update the PCA model to adequately represent the current normal state of the system. A moving window (MW) methodology is introduced to achieve this. MW-PCA operates by firstly learning a model of normal operation on an initial window of data. Then, as each new normal data sample is received it is included in the data window at the expense of the oldest data sample, and a new PCA model computed [25]. The T^2 and Q statistics and their respective confidence limits (T^2_α and Q_α) are also updated for evaluation of the next new sample point.

If $\boldsymbol{F}_i$ denotes the moving window matrix of frequency samples and w is the selected window size, then at the i-th sample instant ($i \geq w$) $\boldsymbol{F}_i$ is simply given by:

$$\boldsymbol{F}_i^T = [\ \ \boldsymbol{f}_{i-w+1}^T \quad \cdots \quad \boldsymbol{f}_{i-1}^T \quad \boldsymbol{f}_i^T \ \ ] \tag{5}$$

and the associated k component PCA model score and loading matrices denoted as, $\boldsymbol{T}^{(i)}$ and $\boldsymbol{P}^{(i)}$, respectively. The, T^2 and Q statistics are given by equations 2 and 3 with $\boldsymbol{P} = \boldsymbol{P}^{(i)}$, with respective confidence intervals denoted as $T^{2(i)}_\alpha$ and $Q^{(i)}_\alpha$.

TABLE I: Consecutive violations against false alarm rate

No. Consecutive warnings used (for 99% confidence limit)	Q Statistic False Alarm Rate
1	16.04%
2	14.58%
3	9.16%
4	3.7%
5	3.4%
6	3.4%

C. Multiple event detection and classification

Due to the high correlation between frequency variables in a power system during normal operation only the first principal component, $\boldsymbol{p}_1$ is required to construct the PCA model. Our previous work, [8], shows that $\boldsymbol{p}_1$ captures more than 99% of the total variance in the frequency data collected from the UK and Irish Power System, and thus captures significant system variation.

As already noted, to construct the MEDC PCA model frequency data related to normal operation of the power must be used. Hence, for the i-th data window when a new sample is received its T^2 and Q statistics are evaluated (with respect to the i-th data window) and if $T^2 \leq T^2_\alpha$ and $Q \leq Q_\alpha$, the system is deemed to be operating normally and the moving window is updated to include the next sample. Otherwise, if one or both the confidence limits have been violated the data point is excluded as a fault may have occurred in the system. To reduce the potential for false alarms a threshold of 5 or more consecutive violations of the 99% confidence limits was defined as the trigger for the MEDC method. Hence, an event is only deemed to have occurred if the T^2 or Q statistic values exceed their confidence limits for the current and previous 4 samples. The value of 5 was chosen based on an evaluation of the false alarm rate as a function of the consecutive violation threshold level for normal operation data, as presented in TABLE I. This revealed that the false alarm rate stabilized for threshold values of 5 and greater.

If a fault is detected the alarm remains triggered until normal operation is resumed, as determined by the statistics of the frequency samples dropping below their respective confidence limits. Whilst an alarm is triggered the PCA model is held constant at pre-alarm values, after the event is cleared the model continues to update with post event data, to avoid contamination by data for abnormal system behaviour. The underlying assumption is that the system topology returns to normal once a fault has been cleared.

The basic process for the proposed MW-PCA based MEDC methodology is presented in Fig. 1. The methodology involves two main components: 1) an off-line training system; and 2) on-line monitoring system. The off-line training system uses historic frequency data from periods of normal power system operation to construct an initial PCA model for varying window sizes. Off-line training concludes when a PCA model is identified with a false alarm rate of less than 1%. It was

found that for the power system studied in this work a window size of between 3 and 6 hours is able to achieve a false alarm rate of less than 1%. A more extensive study into false alarm rates for varying window size is presented in Section V-B. This initial off-line training process allows the MEDC to be easily adapted to different power systems.

The on-line monitoring system can be broken down into multiple steps, as shown in Fig 1, relying on the combination of the T^2 and Q statistics to determine and detect the occurrence of specific faults in the power system. When both statistics are inside their confidence limits, the system is considered to be in normal operation, and therefore, the T^2 and Q statistics, and their respective confidence limits T^2_α and Q_α, are updated.

If the Q statistic is above its confidence limit, it indicates a significant deviation of variables in the system from each other, and in a power system context this implies the occurrence of an islanding event.

If the T^2 statistic, representing the weighted distance from the target, is above its confidence limit it indicates that the system has deviated from its target (50 Hz for UK mains frequency), and thus the occurrence of a significant generation load mismatch event. Note that the method naturally defines a mismatch event as being a frequency deviation which is greater than the normal frequency variation in the system, and depending on the value of the selected confidence limit (typically 95 or 99%) is a deviation that occurs less than 5 or 1% of the time.

Generation load mismatch events can be further classified into net loss of load events (causing a significant rise in frequency) and net generation dip events (causing a significant drop in frequency). Within the MW-PCA MEDC statistical framework these events can be distinguished under both islanding and not islanding conditions. If the system has not islanded (in which case $Q < Q_\alpha$) all the frequency variables from different locations are considered to be synchronised. Then a single system level classification index can be defined as $C_0 = (\boldsymbol{f} - \bar{\boldsymbol{f}})\boldsymbol{p}_1$ and a high frequency event is deemed to have occurred if

$$C_0 > C_{\alpha H} \tag{6}$$

where $C_{\alpha H} = \sqrt{\lambda_1 T^2_\alpha}$ is the high frequency threshold. Here, λ_1 is the eigenvalue associated with the first principal component of data window $\boldsymbol{F}_i$ and $\bar{\boldsymbol{f}}$ is the mean of the data window. Similarly, a low frequency event is deemed to have occurred if

$$C_0 < C_{\alpha L} \tag{7}$$

where $C_{\alpha L} = -\sqrt{\lambda_1 T^2_\alpha}$ is the low frequency threshold. Derivation of equation (6) and (7) are provided in Appendix A.

When $Q > Q_\alpha$, signifying the occurrence of an islanding event, frequency variables from different locations in the power system cluster into groups corresponding to the different islands. These different groups can be identified by analysis of PCA contribution plots [8]. If the frequency variables are decomposed into subspaces corresponding to the different island groups, the frequency variables in the j-th subspace can be denoted as $f^{(j)} \in \Re^{1\times mj}$, with the corresponding loading

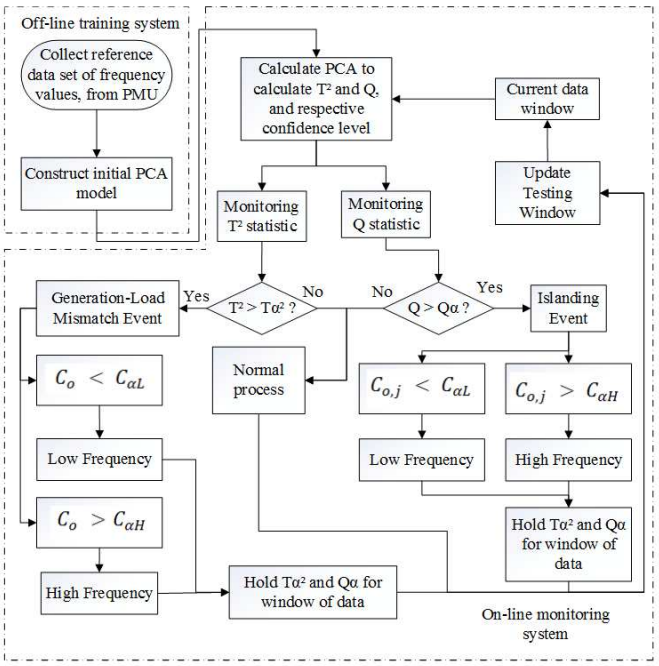


Fig. 1: Flow chart for the proposed MW-PCA MEDC method

vector given by $p_1^{(j)} \in \Re^{mi\times 1}$. High and low frequency events can then be detected for the j-th island if

$$C_{0,j} > C_{\alpha H} \tag{8}$$

and

$$C_{0,j} < C_{\alpha L} \tag{9}$$

respectively, where the j-th subspace classification index is computed as $C_{0,j} = (\boldsymbol{f}^{(j)} - \bar{\boldsymbol{f}}^{(j)})\boldsymbol{p}_1^{(j)}$.

An important consideration with MW-PCA based MEDC is the tuning of the window size as it presents a trade-off between computational efficiency and the ability to accurately detect events [25]. A smaller window size, using less data, is less robust than a larger window as it may adapt to changes in the system too quickly leading to the misinterpretation of normal behaviour in the power system as a fault in the power system. On the other hand, a larger window size will classify everything inside it as normal operation, averaging out larger transients and potentially missing gradually developing events. In addition, it may not adequately discount older data that is not fully representative of the current operating state of the time-varying power system.

Thus, for MW-PCA based MEDC to be successful a window size has to be selected that will have the ability to follow trends in frequency variation in the power system, recognise real faults when they occur, but not trigger for transients that are part of normal operation.

III. EVALUATION WITH SIMULATION CASE STUDIES

To demonstrate the capabilities of the proposed MEDC method a number of simulation studies were conducted on the standard IEEE-9 bus test system (Fig.2) using DIgSILENT

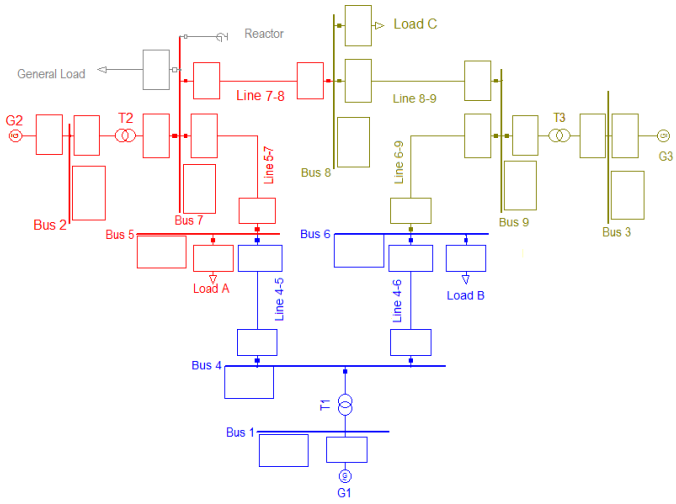


Fig. 2: Configuration of the simulated power system model, in DigSilent PowerFactory.

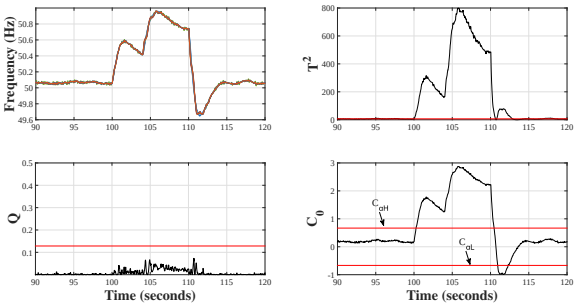


Fig. 3: Zoomed in section of the cascading loss of load events for simulated case study 1, showing the frequency, the T^2 and Q statistics for detection of the loss of loads and the classification index, C_0 for the simulated event)

PowerFactory [26] . This test system consists of 3 generators, 4 loads, 6 transmission lines and 9 buses. A PMU was placed at each bus in the system (giving a total of 9 PMUs) and the sampling rate was set at 100 Hz.

Using this test system three case study scenarios representing different frequency events that can occur on a power system were investigated, namely: (1) loss of load from multiple locations in a cascading fashion; (2) dip in generation and loss of load in a cascading fashion, and; (3) double line trip followed by a cascading line trip leading to multiple island formation.

A. Case 1 - Multiple cascading loss of load

Case study 1 is a cascading double loss of load event, where Load B is lost at 100 seconds followed by Load C 4 seconds later. The frequency plot for a zoomed in section of the simulation is shown in Fig.3, along with a zoomed in view of the MEDC results for detection and classification. It can be observed from the frequency plot that the system frequency begins to rise at 100 seconds to a frequency of 50.6 Hz, due to the loss of Load B. The system then experiences a second rise in frequency to 50.9 Hz at 104 seconds with the loss of Load C. Both loads are then restored to the system at 110 seconds, resulting in a short ($<$ 1 second) under frequency transient before the system becomes stable again.

From the MEDC results in Fig.3, it can be seen that the T^2 statistic detected the high frequency events (Loss of Loads B and C) and the under frequency transient after the loads were restored to the system and that the Q statistic correctly didn't trigger any islanding event occurring in the system.

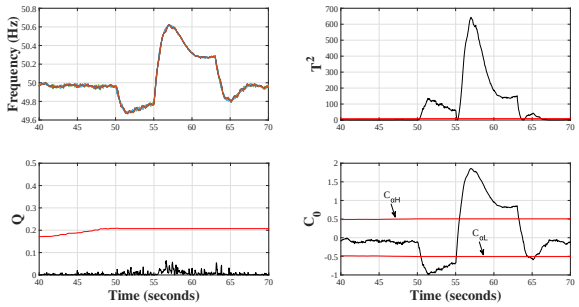


Fig. 4: Zoomed in section of case study 2, showing the dip in generation(generator 1) at 50 seconds, followed by the loss of Load A at 55 seconds. The simulated system frequency, the detection statistics T^2 and Q, and the classification index, C_0 are presented in the plots

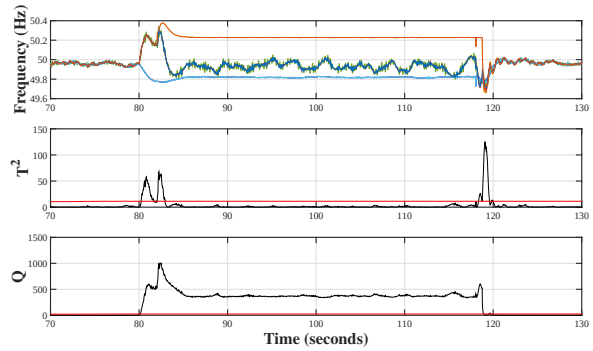
Furthermore, when the T^2 statistic detected a generation load mismatch event between 100 to 112 seconds, C_0 exceeds the upper threshold $C_{\alpha H}$, from 100 seconds to 110 seconds, indicating a high frequency or a net loss of load event experienced on the system. At 111 seconds, C_0 exceeds the lower threshold $C_{\alpha L}$, from 111 s to 112 s, correctly indicating a lower frequency or a net loss of generation event on the system. It should be noted that in general frequency based MEDC methodologies can only identify the 'net' loss of load or generation resulting from a combination of load/generation losses. They cannot resolve individual components occurring simultaneously.

B. Case 2 - Cascading generation dip and loss of load

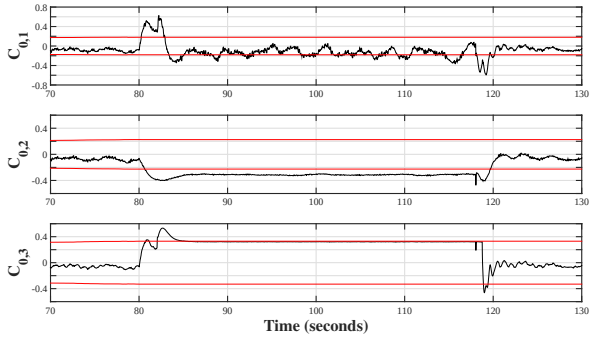
For case study 2, a generation dip was experienced on the system followed by a loss of load a few seconds later. In this case generator 1 experiences a dip in generation at 50 seconds, which is followed by the loss of Load A 5 seconds later. The load and generator are both then restored at 63 seconds. From the frequency plot in Fig.4 it can be noticed that the system frequency drops to around 49.7 Hz after the drop in generation at 50 seconds. The system frequency then experiences a sharp rise to 50.6 Hz after the loss of Load A at 55 seconds before returning to nominal frequency at 64 seconds, following a short under frequency transient due to reconnection.

Fig.4, which shows a zoomed view of the frequency event and MEDC results, demonstrates that MEDC was able to detect the generation dip at 50 seconds with the T^2 statistic violating it's confidence limits followed by the loss of load event at 55 seconds when the confidence limit is violated. Again the short under frequency transient was also successfully detected. The Q statistic also operates correctly by remaining below its confidence limit throughout the event.

In this case when the T^2 statistic detected a generation load match event at 50 seconds the classification index, C_0, exceeds the lower threshold $C_{\alpha L}$ until 55 seconds agreeing with the frequency plot. At 56 seconds a second generation load mismatch event is detected. This time C_0 exceeds the upper threshold $C_{\alpha H}$ from 56 seconds to 63 seconds, indicating



(a) MEDC detection results



(b) MEDC classification results

Fig. 5: A section of the cascading line trip events described in case study 3, showing: (a) the MEDC detection ability, and; (b) the separate classification indexes $C_{0,1} - C_{0,3}$ for the islands in the power system following the line trips

a high frequency or a net loss of load event experienced on the system. Finally at 64 seconds, C_0 exceeds the lower threshold $C_{\alpha L}$ again for 1 second, which corresponds to the short under frequency transient experienced in the system. The classification index was able to classify both the low frequency and high frequency events due to the fact that the loss of load was substantially greater that the loss of generation and hence the event transitioned from a net loss of generation to a net loss of load event.

C. Case 3 - Cascading line trip

Case study 3 involves a double line trip followed by a cascading line loss in the system. Lines 4-5 and 6-9 are lost at 80 seconds causing the system to experience both high and low frequency events simultaneously on different buses in the network as two islands are formed. Two seconds later Line 7-8 is lost, causing an even greater frequency deviation on 3 buses in the system and the formation of a third island. After 38 seconds Lines 4-5 and 6-9 were restored to the system, with Line 7-8 restored cascadingly 0.5 seconds later.

The frequency plot for this case is presented at the top of Fig.5(a), along with MEDC detection results. From the detection results it can be observed that the T^2 statistic was able to detect the frequency deviations in the system corresponding to the simultaneous loss of Lines 4-5 and 6-9, the loss of Line 7-8, and the return of lines to the system. The Q statistic correctly detects the occurrence of an island when Lines 4-5 and 6-9 are lost, and the return to mains

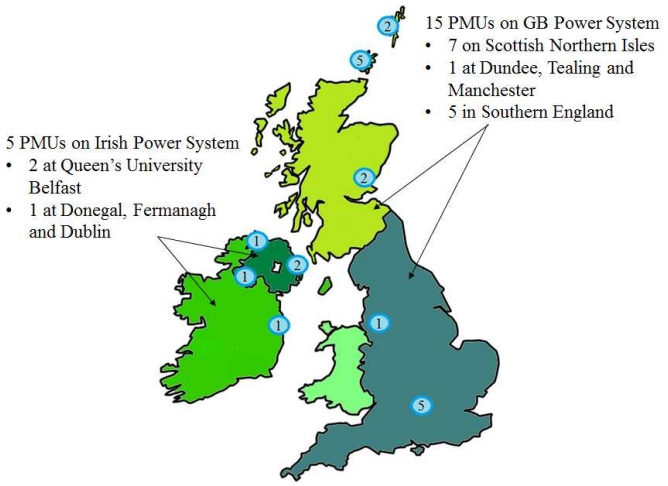


Fig. 6: Location of OpenPMUs in the UK and Irish power system [27]

frequency when all lines are reconnected. It cannot distinguish the simultaneous occurrence of the second island following the loss of Line 7-8, but this can readily be deduced from the principal component contribution plots, as discussed in [8].

When an islanding event occurs each part of the network maintains its own frequency balance, with the result that different islands may experience different local frequency events. For example, island 1 will experience a high frequency event while island 2 experiences a low frequency event if the net flow of power was from island 1 to island 2 immediately prior to the event. Application of the system level classification index, C_0, in not appropriate in these circumstances since it will only provide a single classification which will tend to reflect the more dominant frequency deviation. Instead, by identifying the different islands from the PCA contribution plots, classification indexes can be computed for each islanded system and the local frequency event classified using equations (8) and (9), along with the the upper ($C_{\alpha H}$) and lower ($C_{\alpha L}$) thresholds. Fig. 5(b) shows the classification results using this approach for the three islands that are created by the loss of Lines 4-5, 6-9 and 7-8.

IV. REAL POWER SYSTEM CASE STUDY

To further demonstrate the utility of the MW-PCA based MEDC methodology this section evaluates its performance for real event data collected from the UK power system. The data for analysis was collected using a wide area measuring system consisting of PMUs installed at sites of interest across the UK and Ireland. These PMUs were developed at Queen's University Belfast as part of the OpenPMU project [27] with support from Scottish and Southern Energy Ltd. Fig. 6 shows the locations of the OpenPMUs installed in the UK system. The PMUs record frequency, phase angle and voltage magnitude at each of the sites and report at 10 Hz [27]. The case study frequency data used for analysis was taken from 6 of the sites mentioned above: one from Southern England (f_1), one from Manchester (f_2), and four from the Orkney Islands (f_3,f_4,f_5,f_6). A window size of 6 hours was used with the MW-PCA models.

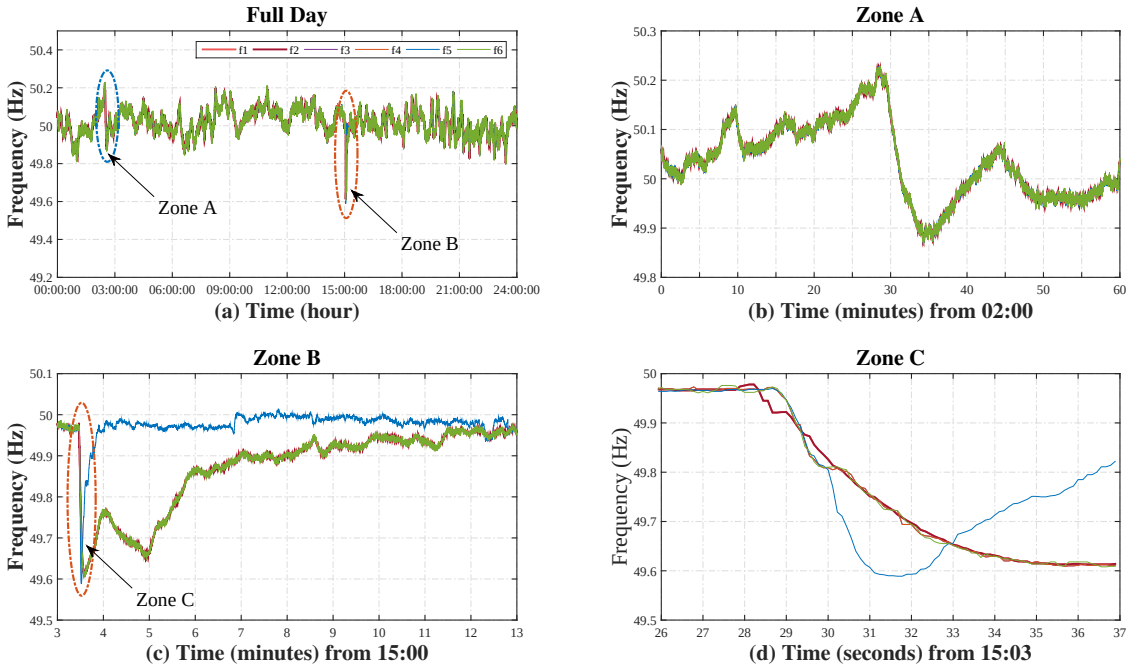


Fig. 7: Frequency plots for Case 1: (a) Frequency plot for 24 hour period (b) recorded high frequency; (c) interconnector trip and island; (d) frequency nadir due to island

A. Case 1: Loss of Load, Interconnector Trip and Islanding Event-30/09/2012

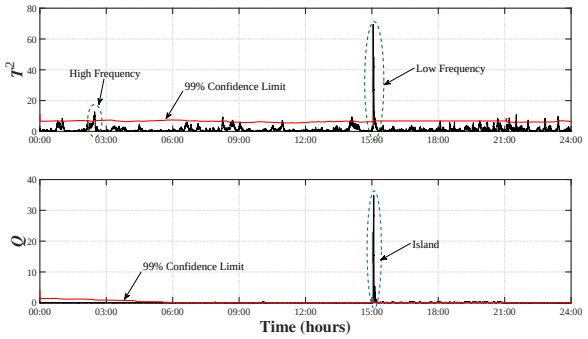
A plot of the recorded frequency values at the 6 PMU sites for case 1 is shown in Fig 7(a), with a magnified view of 3 sections of interest presented in Fig 7(b), (c) and (d).

It can be seen that on the morning of 30th September 2012, there was a high frequency recorded on the system, causing the frequency at all 6 PMU sites to exceed 50.24 Hz, at 02:28. This high frequency event represents a loss of load on the power system.

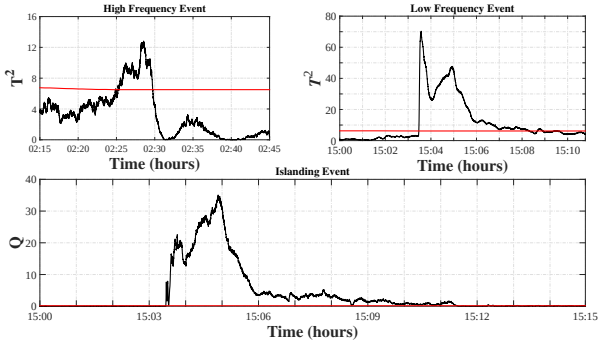
Later on the same day, at 15:03:28, a low frequency event on the power system caused a drop in frequency across the system. This low frequency event was attributed to the occurrence of a fault on the 2GW 'GB-France HVDC Cross-Channel Interconnector', causing the instantaneous loss of 1GW of power to the GB system. It was observed, in Fig 7(d), firstly that the recorded frequency at two PMU sites (f_1 and f_2), deviated slightly from the other four frequency samples.

This is due to the location of the interconnector at Sellindge, in Southern England. After the interconnector trip it takes a certain length of time for frequency in the north of Scotland to be affected. Since the f_1 and f_2 PMUs are geographically closer they see the fault before the North Scotland PMUs.

Also, Fig 7(d) shows that the recorded frequency of the power system fell from 49.97 Hz to 49.6 Hz in the 7 seconds immediately after the loss of the interconnector at 5 of the 6 PMU sites. One site (f_5) dropped at a quicker rate, 49.97 Hz to 49.59 Hz in 3 seconds, resulting in islanding protection operation being triggered at this site, leading to the added loss of embedded generation. After the triggering of the islanding protection the PMU at f_5 recorded the occurrence of an islanding event immediately after the frequency fell below 49.8 Hz which lasting for just over 9 minutes, as shown in Fig. 7(c).



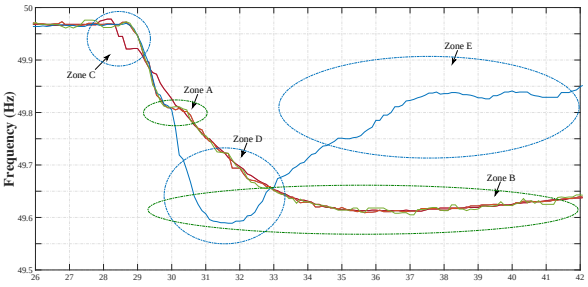
(a) MEDC Results for full day



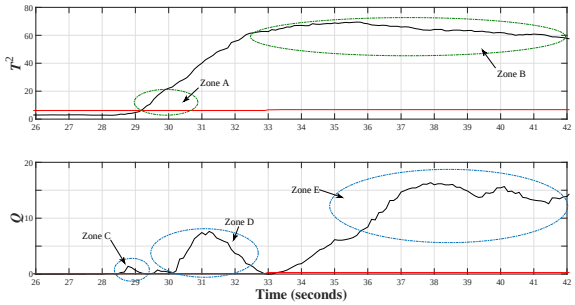
(b) Zoomed in MEDC Results

Fig. 8: MEDC monitoring results for Case 1 showing the detection of the three events mentioned previously. The high frequency at 02:28:03, while the interconnector trip and islanding events detected at 15:03:30

The results from the MEDC monitoring method for case 1 for the full day are illustrated in Fig 8. The confidence



(a) Frequency Plot



(b) MEDI monitoring results

Fig. 9: Magnified view of: (a) frequency plot; (b) MEDC monitoring results for case 1 from 15:03:25 to 15:03:42. Showing: Zone A-frequency dropping below MEDC threshold, Zone B-frequency nadir, Zone C-measurement delay, Zone D-initial island at f_5 and Zone E-main island at f_5

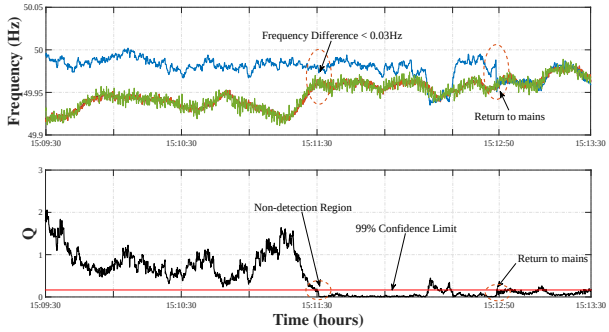


Fig. 10: Magnified view of frequency (upper plot) and Q statistic of MEDC method (lower plot) for conclusion of islanding event for case 1 from 15:09:30 to 15:13:30, showing return to mains frequency

threshold was set at 99% to avoid excessive false alarms being triggered. As observed the method has the ability to detect the three events highlighted previously.

1) High Frequency: From Fig 8 it can be seen that the high frequency was detected by the T^2 statistic, corresponding to the frequency plot in Fig 7(b). This high frequency was detected when the frequency exceeded 50.1 Hz.

2) Low Frequency and Island: Fig 9 shows a magnified section of both the frequency plot and MEDC results from 15:03:25 to 15:03:42, around the time of the previously mentioned inter-connector trip. It can be clearly seen that the T^2 statistic (upper plot, Fig. 9(b)) detected a substantial frequency deviation from the target frequency of 50 Hz at 15:03:29 (Zone A). This agrees with the frequency plot in Fig. 9(a) where the frequency recorded at all sites reaches 49.8 Hz by 15:03:29 before falling to 49.6 Hz by 15:03:35 (Zone B).

At 15:03:28.5 the Q statistic detects the deviation of sites f_1 and f_2 from the other PMU sites (Zone C), agreeing with the frequency plot in Fig. 9(a). This is due to the geographical location of the event on the system causing a delay of measurement at the PMU sites located further North in the Orkney Islands and is therefore a false alarm. This is followed by the frequency f_5 deviating significantly from the others (Zone D) between 30 and 33 seconds, before the main island event (Zone E) is detected by the violation of the calculated Q statistic from its 99% confidence limit at 15:03:34, as shown in Fig. 9(b).

Fig. 10 presents a magnified view of the frequency and Fig. 10 MEDC monitoring result for the Q statistic for the interval where the islanded site returns to mains frequency. It should be noted that with the confidence limit set at 99% it was not possible to distinguish a frequency deviation of less than 0.03 Hz in the system; this can be seen at 15:11:30 in the lower plot in Fig. 10 where the calculated Q statistic does not detect the island for a short time before return to mains at 15:12:50.

B. Case 2: Short Circuit Fault and Islanding Event- 11/11/2014

On the 11th October 2014 there was a low frequency event recorded by the OpenPMU network at 4 of the 6 reference sites, f_3 to f_6, which are located on the Orkney Islands. The other 2 PMU sites, located on the GB mainland, didn't detect any major frequency deviation, and thus this low frequency event is classed as a local event to the North of Scotland. The frequency plot in Fig. 11(a) shows the low frequency event recorded at the 4 Orkney sites occurred at 06:28:46, with an islanding event occurring at f_5 almost instantaneously after the loss of generation, lasting until 07:54:08. Magnified views of the low frequency event as well as the conclusion of the island are illustrated in Fig. 11(b) and (c), respectively.

From this magnified view of the frequency plot Fig. 11(b) it can be seen that the recorded frequency at f_3, f_4 and f_6 deviates slightly from the mains frequency of the system at that time, from 49.93 Hz to 49.84 Hz, before increasing to 50 Hz, both of which are above the threshold of generation load mismatch detection. However, it can be observed that the frequency at f_5 decreases from 49.93 Hz to 49.76 Hz in 0.5 seconds (at 06:28:46) before increasing to 50.19 Hz in a further 0.5 seconds before again dropping to 49.3 Hz in 2 seconds (at 06:28:49), which means that it has gone beyond its detection threshold, and, again similar to case 1, has triggered the islanding protection at that site. From Fig. 11 (c) it can be concluded that the islanding event lasted for just under 90 minutes, with a return to mains frequency recorded as occurring at 07:54:08.

These events were caused by a short circuit fault in the north of Scotland resulting in a loss of generation, leading to the low frequency and islanding event observed at f_5. These events were again used to examine the effectiveness of the proposed MW-PCA MEDC method.

As shown in Fig. 12, the T^2 statistic successfully detected the low frequency at 06:28:46, when the frequency dropped

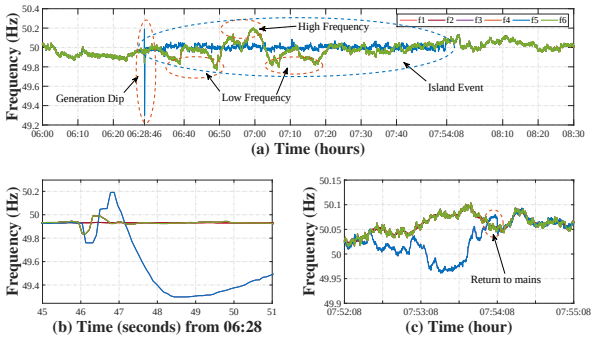


Fig. 11: Frequency plots for Case 2 showing: (a) full event history with low frequency (blue dashed line) and islanding event (green dashed line); (b) low frequency due to short circuit fault and; (c) conclusion of island

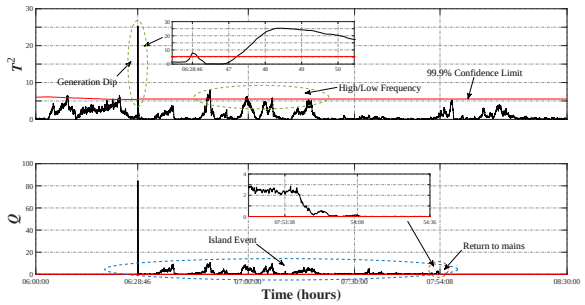


Fig. 12: MEDC monitoring results for case 2 on 11/11/2014. Showing the island (lower plot) and low frequency as well as other generation load mismatch events in the system (upper plot)

to 49.76 Hz, as well as the 3 other instances of the frequency deviating from the target frequency; at 06:49:14 and 07:06:46 the frequency dropped to below 49.8 Hz (MEDC triggered the fault at 49.76 Hz) and at 06:59:37 a high frequency was detected, these detections are consistent with the frequency plot in Fig. 11(a). The Q statistic displayed in the lower plot of Fig. 12 shows successful detection of the islanding event at 06:28:47. Through the Q statistic the MEDC monitoring system also identifies return to mains frequency as occurring at 07:54:08, hence the island event lasted approximately 85 minutes. Again these results agree with the frequency plots in Fig. 11.

V. Discussion

A. Comparison with PCA

To benchmark the performance of MW-PCA based MEDC against traditional PCA [8] the case studies presented in the previous section were used along with a sample of other frequency events that were captured by the OpenPMU network on the UK power system between 2012 and 2015. Overall a sample of 7 islanding (I) events and 15 generation load mismatch (GLMM) events were used in a comparative study. TABLE II provides a summary of the performance of both traditional PCA and MW-PCA based MEDC for fault detection and fault classification for the aforementioned events. The performance metric reported is the number of correct detections/classifications expressed as a percentage of the total number of events.

TABLE II: COMPARISON OF PCA AND MW-PCA based MEDC

Comparison		PCA	MW-PCA based MEDC
Fault Detection	GLMM	100%	100%
	I	100%	100%
Classification	HF	N/A	100%
	LF	N/A	100%
	I	100%	100%

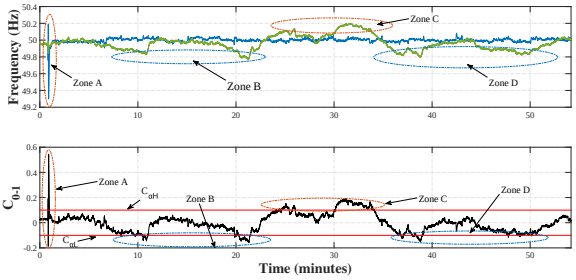


Fig. 13: Section of the fault isolation results from 06:27:56 for case 2. Showing: Zone A-initial frequency drop leading to island, Zone B-low frequency on rest of system during the island, Zone C- high frequency on rest of system during the island, Zone D-low frequency event on rest of system during the island

1) Fault Detection Ability: From TABLE II it can be observed that both the traditional PCA and MW-PCA based algorithms had the ability to detect I events (using the Q statistic) and both were able to detect all of the GLMM events (using the T^2 statistic) for all the events tested.

In contrast to the traditional PCA approach [8], as shown in Fig 8, the MW-PCA method has the ability to adapt to the latest normal operating data as it becomes available, with the control limits varying as the system varies throughout the day. This allows the time-varying characteristics of the power system to be taken into consideration and facilitates real-time updating of the PCA model, leading to a greater situational awareness of the varying power system.

2) Classification Ability: One advantage of the MEDC method over the traditional PCA method is its ability to distinguish between the GLMM events. Again TABLE II is used to compare the classification ability of both methods. Since GLMM classification is not possible with traditional PCA, N/A is recorded as its classification performance. In contrast, using MW-PCA MEDC all of the GLMM events were correctly categorized as either high frequency (HF) or low frequency (LF) events using the classification index derived from the T^2 statistic.

The default classification index, C_0, detects if there is a net deficit or excess in generation for the overall system. However, as discussed in the section III-C, when an islanding event occurs, frequency events are localized to each island and can only be correctly classified through the use of localized classification indexes. If C_0 is inappropriately used it tends to reflect the more dominant frequency deviation. One such successful event classification is shown in Fig. 13, which shows the system high and low frequency being successfully classified, using C_0, for real case study 2.

TABLE III: COMPARISON OF WINDOW SIZE FOR CASE 1

		Reference time of event (defined from a threshold of 0.2 Hz)	Window Size					
			15 m	1 h	3 h	6 h	12 h	24 h
Fault Detection (seconds)	High Frequency	02:28:03	+5	-6	-169	-170	-172	-172
	Low Frequency	15:03:29	0	0	0	0	0	0
	Island	15:03:28	0	0	0	0	0	0
Return to Mains Detection (seconds)	High Frequency	02:29:28	-48	-35	+8	+15	+19	+21
	Low Frequency	15:05:40	+172	+155	+149	+82	+58	+6
	Island	15:12:50	+676	+303	+234	+1	0	0
False Alarm Rate (%)		Q	0.58	0.57	0.55	0.57	0.51	0.48
		T^2	99.9	13.9	0.7	0.17	0.014	0.008
Average Computation Time (seconds)			0.0127	0.211	0.375	0.625	1.133	2.172

A detailed inspection of individual PMU signals would show the direction of frequency change and allow event classification. However, if there are a large number of PMUs to monitor this becomes cumbersome. Furthermore, although the raw frequency data provides a direct visualization of frequency variation, a threshold needs to be defined in order to provide automatic detection of abnormal deviations. Using the C_0 statistic in the MW-PCA MEDC methodology allows this detection and classification to be performed automatically in the low dimension PCA space irrespective of the number of PMUs on the system using an automatically computed event classification threshold. An additional advantage that the PCA based approach offers is that, by fusing data from multiple sites, it enhances the signal to noise ratio and hence enables more robust event detection and classification.

B. Window Size

As discussed in Section II-C an important consideration when using MW-PCA is fine tuning the window size. A number of factors need to be considered in terms of determining the optimal window size, namely, the response time of both the start of detection (fault detection) and end of detection (return to mains), false alarm rate and computational time. To illustrate the impact of window size on performance a comparative study for case 1 of the real data case study (Section IV-A) is presented in TABLE III. All simulations were carried out in Matlab 8.4 (R2014b) on a 3rd generation Intel Core i5 processor with 8GB RAM.

The information presented in TABLE III outlines:

1) the time (in seconds) for the MW-PCA MEDC to detect a fault following its occurrence (Fault Detection),
2) the time (in seconds) for the MW-PCA MEDC to detect a return to mains frequency after a fault has cleared (Return to Mains Detection),
3) the false alarm rate for both the T^2 and Q statistics for each window size,
4) the average computation time (in seconds) to update the PCA model and classify the new sample.

It should be noted that due to the absence of ground truth information the precise onset times of events is not known. Consequently, the reference time for event onsets, as used in TABLE III, is defined as the time at which the frequency first deviates by more than 0.2 Hz from the nominal frequency. The results show that MW-PCA can sometimes detect the fault before this time (hence the negative detection times). This is simply a reflection of the fact that the statistically defined confidence limits of MW-PCA MEDC are more sensitive to abnormal frequency changes. Therefore for slowly changing frequency deviations, such as the high frequency event in Case 2, detection occurs at a lower frequency deviation. For example, in the case of the 12 hour moving window, the high frequency event was detected 172 seconds prior to the reference event onset time, when the frequency deviation was 0.17 Hz.

From the results it can be observed that all of the window sizes evaluated were able to successfully detect the occurrence of the islanding event within a few seconds of onset; however only the larger window sizes provided satisfactory response times for detection of the subsequent return to mains. The low frequency event was also detected in a timely fashion using all window sizes, however again the return to mains frequency detection was poor at smaller window sizes and improved as window size was increased. Notably, the high frequency event was detected most accurately when using a 15 minute window size; but this yielded the worst return to mains detection accuracy of the window sizes tested.

Finally, as expected the false alarm rate decreases with increasing window size however this increases the computational time. Choosing a window size less than 3 hours gave too high a false alarm rate, whilst a window size above 6 hours was considered too computationally inefficient.

C. Computational Complexity

For moderate numbers of PMUs (e.g. a few 100) computational complexity is not a major issue for MW-PCA as only the first PC needs to be computed and this can be done efficiently with a variety of algorithms. For example, for $\boldsymbol{F} \in \Re^{w \times m}$, where w is the window size and m is the number of PMUs, the NIPALS algorithm [28] has $O(wm)$ complexity for estimating the largest PC. Hence, computational complexity scales linearly with both window size and number of PMUs considered.

When $w >> m$, it can be more efficient for MW-PCA model updates to precompute the initial covariance matrix $\boldsymbol{F}^T\boldsymbol{F}$ offline, which has $O(m^2 w)$ complexity. Online updating

of the covariance matrix can then be performed with $O(m^2)$ complexity. This then allows the largest PC to be determined using NIPALS in $O(m^2)$ complexity, with the result that the sliding window PCA model update is independent of window size. For larger window sizes, the computational overhead of computing the initial covariance matrix can be reduced by approximating it using a random subsample of the w data samples over the data window.

The computational cost for online updating of the PCA model becomes a major concern when thousands of PMUs are installed and the data from thousands of locations needs to be analysed simultaneously. In these circumstances a different strategy is required, e.g. employing a distributed implementation and using a multi-block approach to group multiple PMUs into sub-blocks.

VI. Conclusions

This paper proposes a methodology for the detection and classification of multiple events in an electrical power system in real-time, using a moving window based principal component analysis method for wide-area synchronized frequency measurements obtained from a network of PMUs. Results are presented evaluating the methodology on simulated case studies and real events recorded from the UK power system.

From the case studies presented it has been shown that, by thresholding on the T^2 and Q statistics derived from a PCA model of the data, it is possible to discriminate between islanding events and generation load mismatch events, such as line trips, inter-connector trips, generation dip and loss of load events. The proposed MW-PCA based MEDC methodology also has the ability to distinguish between the net loss of load and net loss of generation events in the power system and the ability to automatically define the frequency at which a high or low frequency event occurs for the time varying power system.

Furthermore, the moving window approach adopted for computation of the PCA model allows the time-varying nature of the power system to be taken into consideration, that is, the model is able to track the current normal operating state of the system, thereby improving overall situational awareness of the power system.

The case studies presented also highlight some limitations of the proposed methodology. In particular, it is not able to disaggregate multiple loss of load and generation events occurring in the system simultaneously. It can only detect and classify the consequent net over/under generation. However, this limitation is due to frequency being a global power system variable and is thus a limitation of all frequency based power system analysis techniques. In addition, if the frequency at the islanded site is well matched to the mains frequency ($<$0.03 Hz), then an islanding event may go undetected, but this blind spot is common to all frequency based islanding detection techniques.

Future work will look at a number of enhancements to the MW-PCA based MEDC methodology to address these limitations. Firstly, to improve classification accuracy and the ability to disaggregate multiple events in the power system, a multiple PCA model framework will be developed to enable power system operation during events to be modelled separately from normal operation. Secondly, to improve robustness in terms of detection of islanding events where there is little frequency drift between the island and the rest of the power system voltage and phase angle information will be incorporated into the PCA models.

Appendix A
Derivation of the Classification Index

From Equation (2) the T^2 statistic is given by

$$T^2 = (\boldsymbol{f} - \bar{\boldsymbol{f}})\boldsymbol{P}\Lambda^{-1}\boldsymbol{P}^T(\boldsymbol{f} - \bar{\boldsymbol{f}})^T$$

and a generation load mismatch event is deemed to have occurred on the power system when [8]:

$$T^2 = (\boldsymbol{f} - \bar{\boldsymbol{f}})\boldsymbol{P}\Lambda^{-1}\boldsymbol{P}^T(\boldsymbol{f} - \bar{\boldsymbol{f}})^T > T_\alpha^2 \tag{10}$$

If only 1 principal component is retained, equation (10) reduces to

$$(\boldsymbol{f} - \bar{\boldsymbol{f}})\boldsymbol{p}_1\lambda_1^{-1}\boldsymbol{p}_1^T(\boldsymbol{f} - \bar{\boldsymbol{f}})^T > T_\alpha^2 \tag{11}$$

which in turn can be expressed as

$$\lambda_1^{-1}||(\boldsymbol{f} - \bar{\boldsymbol{f}})\boldsymbol{p}_1||^2 > T_\alpha^2$$

Finally, this can be rearranged to yield the relationship

$$(\boldsymbol{f} - \bar{\boldsymbol{f}})\boldsymbol{p}_1 > \sqrt{\lambda_1 T_\alpha^2}$$

If this condition is satisfied it implies that the frequency deviation from its target exceeds the positive threshold, indicating a high frequency event or net loss of load event has occurred. Similarly if

$$(\boldsymbol{f} - \bar{\boldsymbol{f}})\boldsymbol{p}_1 < -\sqrt{\lambda_1 T_\alpha^2}$$

is true it indicates that the frequency deviation exceeds the negative threshold and corresponds to a low frequency or net loss of generation event.

Acknowledgment

The authors would like to acknowledge Scottish & Southern Energy Power Distribution Ltd. for their assistance in sharing of UK power system data. The authors would also like to acknowledge the Department of Education and Learning, Northern Ireland for their funding of the work carried out in this paper.

References

[1] D. Novosel, K. Vu, V. Centeno, S. Skok, and M. Begovic, "Benefits of synchronized-measurement technology for power-grid applications," in *Proc. 40th Annu. Hawaii Int. Conf. System Sciences*, 2007, pp. 118–118.

[2] S. Abraham, H. Dhaliwal, R. J. Efford, L. J. Keen, A. McLellan, J. Manley, K. Vollman, N. J. Diaz, T. Ridge *et al.*, *Final report on the August 14, 2003 blackout in the United states and Canada: causes and recommendations*. US-Canada Power System Outage Task Force, 2004.

[3] E. Barocio, B. C. Pal, D. Fabozzi, and N. F. Thornhill, "Detection and visualization of power system disturbances using principal component analysis," in *Proc. Symposium on Bulk Power System Dynamics and Control - IX Optimization, Security and Control of the Emerging Power Grid (IREP)*, 2013, pp. 1–10.

[4] W. Liu, Z. Lin, F. Wen, and G. Ledwich, "A wide area monitoring system based load restoration method," *IEEE Trans. on Power Syst.*, vol. 28, no. 2, pp. 2025–2034, 2013.

[5] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2002.

[6] E. L. Russell, L. H. Chiang, and R. D. Braatz, *Data-driven methods for fault detection and diagnosis in chemical processes*. Springer Science & Business Media, 2012.

[7] S. Meher and P. Maben, "Face recognition and facial expression identification using PCA," in *2014 IEEE International Advance Computing Conference (IACC)*, Feb 2014, pp. 1093–1098.

[8] X. Liu, D. Laverty, R. Best, K. Li, D. Morrow, and S. McLoone, "Principal component analysis of wide-area phasor measurements for islanding detection-a geometric view," *IEEE Trans. on Power Del.*, vol. 30, no. 2, pp. 976–985, 2015.

[9] A. Sinha and K. K. Chowdoju, "Power system fault detection classification based on PCA and PNN," in *2011 International Conference on Emerging Trends in Electrical and Computer Technology (ICETECT)*, 2011, pp. 111–115.

[10] B. Mnassri, M. El Adel, and M. Ouladsine, "Fault localization using principal component analysis based on a new contribution to the squared prediction error," in *2008 16th Mediterranean Conference on Control and Automation*, 2008, pp. 65–70.

[11] Z. Zhong, C. Xu, B. J. Billian, L. Zhang, S.-J. S. Tsai, R. W. Conners, V. Centeno, A. G. Phadke, Y. Liu *et al.*, "Power system frequency monitoring network (FNET) implementation," *IEEE Trans. on Power Syst.*, vol. 20, no. 4, pp. 1914–1921, 2005.

[12] Z. Lin, T. Xia, Y. Ye, Y. Zhang, L. Chen, Y. Liu, K. Tomsovic, T. Bilke, and F. Wen, "Application of wide area measurement systems to islanding detection of bulk power systems," *IEEE Trans. on Power Syst.*, vol. 28, no. 2, pp. 2006–2015, 2013.

[13] J. Dong, J. Zuo, L. Wang, K. S. Kook, I.-Y. Chung, Y. Liu, S. Affare, B. Rogers, and M. Ingram, "Analysis of power system disturbances based on wide-area frequency measurements," in *Proc. 2007 IEEE Power Engineering Society General Meeting*, 2007, pp. 1–8.

[14] W. Wang, L. He, P. Markham, H. Qi, Y. Liu, Q. C. Cao, and L. M. Tolbert, "Multiple event detection and recognition through sparse unmixing for high-resolution situational awareness in power grid," *IEEE Trans. on Smart Grid*, vol. 5, no. 4, pp. 1654–1664, 2014.

[15] A. Bykhovsky and J. H. Chow, "Power system disturbance identification from recorded dynamic data at the northfield substation," *International journal of Electrical Power & Energy Systems*, vol. 25, no. 10, pp. 787–795, 2003.

[16] Y. Guo, K. Li, D. Laverty, and Y. Xue, "Synchrophasor-based islanding detection for distributed generation systems using systematic principal component analysis approaches," *IEEE Trans. on Power Del.*, vol. 30, no. 6, pp. 2544–2552, 2015.

[17] G. Gajjar and S. Soman, "Auto detection of power system events using wide area frequency measurements," in *2014 Eighteenth Nationa Power Systems Conference (NPSC)*, 2014, pp. 1–6.

[18] L. Xie, Y. Chen, and P. R. Kumar, "Dimensionality reduction of synchrophasor data for early event detection: Linearized analysis," *IEEE Trans. on Power Syst.*, vol. 29, no. 6, pp. 2784–2794, 2014.

[19] D. Phillips and T. Overbye, "Distribution system event detection and classification using local voltage measurements," in *2014 Power and Energy Conference at Illinois (PECI)*, 2014, pp. 1–4.

[20] A. Gaouda, S. Kanoun, M. Salama, A. Chikhani *et al.*, "Pattern recognition applications for power system disturbance classification," *IEEE Trans. on Power Del.*, vol. 17, no. 3, pp. 677–683, 2002.

[21] G. Zheng and R. Craven, "Multiclass support vector machines for power system disturbances classification based on wide-area frequency measurements," in *2011 Proceedings of IEEE Southeastcon*, 2011, pp. 68–72.

[22] R. M. Gardner and Y. Liu, "Generation-load mismatch detection and analysis," *IEEE Trans. on Smart Grid*, vol. 3, no. 1, pp. 105–112, 2012.

[23] L. H. Chiang, R. D. Braatz, and E. L. Russell, *Fault detection and diagnosis in industrial systems*. Springer Science & Business Media, 2001.

[24] J. E. Jackson, *A user's guide to principal components*. John Wiley & Sons, 2005, vol. 587.

[25] X. Wang, U. Kruger, and G. W. Irwin, "Process monitoring approach using fast moving window PCA," *Industrial & Engineering Chemistry Research*, vol. 44, no. 15, pp. 5691–5702, 2005.

[26] Digsilent. [Online]. Available: http://www.digsilent.de/

[27] D. M. Laverty, R. J. Best, P. Brogan, I. Al Khatib, L. Vanfretti, and D. J. Morrow, "The OpenPMU platform for open-source phasor measurements," *IEEE Trans. on Instrumentation and Measurement*, vol. 62, no. 4, pp. 701–709, 2013.

[28] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and intelligent laboratory systems*, vol. 2, no. 1-3, pp. 37–52, 1987.